

Distinguishing multi-partite states by local measurements

Cécilia Lancien^{1,2} and Andreas Winter^{2,3}

¹*Centre de Mathématiques Laurent Schwartz,
École Polytechnique, 91128 Palaiseau Cedex, France**

²*Department of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.*

³*Centre for Quantum Technologies, National University of Singapore,
2 Science Drive 3, Singapore 117542, Singapore[†]*

(Dated: 14 June 2012)

We analyze the distinguishability norm on the states of a multi-partite system, defined by local measurements. Concretely, we show that the norm associated to a tensor product of sufficiently symmetric measurements is essentially equivalent to a multi-partite generalisation of the non-commutative ℓ_2 -norm (aka Hilbert-Schmidt norm): in comparing the two, the constants of domination depend only on the number of parties but not on the Hilbert spaces dimensions.

We discuss implications of this result on the corresponding norms for the class of all measurements implementable by local operations and classical communication (LOCC), and in particular on the leading order optimality of multi-party data hiding schemes.

I. DISTINGUISHABILITY NORMS

The task of distinguishing quantum states from accessible experimental data is at the heart of quantum information theory, appearing right at its historical beginnings – see [9, 10], and [13] for general reference. Indeed, the special case on which we are focussing in this paper, the discrimination of two states, is the generalisation of hypothesis testing in classical statistics. There, the optimal discrimination between two hypotheses, modelled as (for simplicity: discrete) probability distributions P_0 and P_1 , with prior probabilities q and $1 - q$, respectively, is given by the maximum likelihood rule [6]. The minimum error probability is thus given by

$$\Pr\{\text{error}\} = \frac{1}{2}(1 - \|qP_0 - (1 - q)P_1\|_1),$$

with the usual ℓ_1 -norm $\|\Delta\|_1 = \sum_{x \in \mathcal{X}} |\Delta_x|$.

In this paper, we shall denote by the same symbol its non-commutative generalisation $\|\Delta\|_1 = \text{Tr} |\Delta|$, i.e. the sum of the singular values of Δ , also known as trace norm.

Owing to the particular role played by *measurement* in quantum mechanics, however, any restriction on the set of available measurements leads to a specific norm on density operators: any decision in the discrimination task must be based on measurement results. Specifically, let the two hypotheses be two quantum states (density operators) ρ_0 and ρ_1 on some Hilbert space \mathcal{H} , with prior probabilities q and $1 - q$, respectively. A generic measurement M , i.e. a positive operator valued measure (POVM, aka partition of unity), is given by positive semidefinite operators

$$M_x \geq 0, \quad \text{s.t.} \quad \sum_{x \in \mathcal{X}} M_x = \mathbb{1}.$$

(In this paper, POVMs will generally be discrete and Hilbert spaces will always be of finite dimension. With suitable adaptations to the proofs, however, our results carry over to general POVMs)

* cecilia.lancien@polytechnique.edu

† a.j.winter@bris.ac.uk

and infinite dimension.) The Born rule for measurements postulates that the state ρ_i generates a distribution P_i on the outputs of the measurement:

$$P_i(x) = \text{Tr } \rho_i M_x,$$

and hence the minimum error probability in any decision based on i is

$$\begin{aligned} \Pr\{\text{error}\} &= \frac{1}{2} (1 - \|qP_0 - (1-q)P_1\|_1) \\ &= \frac{1}{2} \left(1 - \sum_{x \in \mathcal{X}} |\text{Tr}(q\rho_0 - (1-q)\rho_1)M_x| \right) \\ &=: \frac{1}{2} (1 - \|q\rho_0 - (1-q)\rho_1\|_M). \end{aligned}$$

Observe that $\|\Delta\|_M = \sum_{x \in \mathcal{X}} |\text{Tr } \Delta M_x|$ is a seminorm: it is non-negative, homogeneous and obeys the triangle inequality. However, it may vanish on $\Delta \neq 0$. This is excluded if the measurement M is *informationally complete*, meaning that the operators M_x span all the operators over the Hilbert space: $\text{span}\{M_x : x \in \mathcal{X}\} = \mathcal{B}(\mathcal{H})$.

If not one but a whole set \mathbf{M} of measurements is given, from which the experimenter may choose, we have an equally natural (semi-)norm

$$\|\Delta\|_{\mathbf{M}} = \sup_{M \in \mathbf{M}} \|\Delta\|_M,$$

in terms of which the minimum error probability is expressed as $\frac{1}{2} (1 - \|q\rho_0 - (1-q)\rho_1\|_{\mathbf{M}})$. These norms, under certain restrictions of interest on the measurement, will be the object of study in the present paper, and in particular their comparison with the trace norm, which by a classic observation of Holevo [10] and Helstrom [9] equals the distinguishability norm under the set of all possible measurements:

$$\|\Delta\|_{\mathbf{ALL}} = \sup_{M \text{ any POVM}} \|\Delta\|_M = \|\Delta\|_1 = \text{Tr } |\Delta|.$$

In this spirit, we continue an investigation begun in [12], addressing some of the questions left open there. The reader is referred to that paper for further information about distinguishability norms and their interpretation in terms of the geometry of certain convex bodies of operators. Note however that many results from [12] are restricted to traceless operators $\Delta = \frac{1}{2}(\rho_0 - \rho_1)$, corresponding to equal prior probabilities $q = 1 - q = \frac{1}{2}$. Of course, mathematically and also in view of applications with unequal prior probabilities, it makes sense to lift this restriction.

The structure of the rest of the paper is as follows: In section II we define the measurements and some classes of measurements we will be interested in, introducing also a multi-partite generalisation of the non-commutative ℓ_2 -norm (aka Hilbert-Schmidt norm), denoted $\|\cdot\|_{2(K)}$. In section III we then state and prove our main results comparing measurement norms with 2-norms, while in section IV we move on to relations with the trace norm and the application of our results to so-called *data hiding*. We conclude in section V with a brief discussion. Appendix A is devoted to the technical parts of the proof of the main result, building on ideas from [1, 12].

II. PROJECTIVE t -DESIGNS AND LOCC MEASUREMENTS

As explained in the introduction, if we want to use a single measurement to define a norm it has to be *informationally complete*. Among those, there are measurements with special symmetry properties known as (*projective*) *designs* – see [7, 11, 17] and [1].

Definition 1 A rank-one POVM $M = (M_x)_{x \in \mathcal{X}}$ on a d -dimensional Hilbert space \mathcal{H} is called a t -design if the ensemble $\{p_x, P_x\}$ of rank-one projectors, with $p_x = \frac{1}{d} \text{Tr } M_x$ and $P_x = \frac{M_x}{\text{Tr } M_x}$, is a projective (weighted) t -design in the usual sense [7, 11, 17], i.e. if

$$\sum_{x \in \mathcal{X}} p_x P_x^{\otimes t} = \int d\psi |\psi\rangle\langle\psi|^{\otimes t}.$$

where the integral is over the uniform (unitary invariant) probability measure on the pure states of \mathcal{H} .

Note that

$$\int d\psi |\psi\rangle\langle\psi|^{\otimes t} = \frac{1}{\binom{d+t-1}{t}} \Pi_{\text{Sym}} = \frac{1}{d(d+1) \cdots (d+t-1)} \sum_{\pi \in \mathfrak{S}_t} U_\pi,$$

where Π_{Sym} is the projector onto the completely symmetric subspace of $\mathcal{H}^{\otimes t}$, i.e. the subspace of $\mathcal{H}^{\otimes t}$ invariant under all the permutation unitaries $U_\pi |v_1\rangle \otimes \cdots \otimes |v_t\rangle = |v_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |v_{\pi^{-1}(t)}\rangle$.

We shall be concerned with multi-partite quantum systems. To fix notation for the rest of the paper, let $\mathcal{H}_1, \dots, \mathcal{H}_K$ be K finite dimensional Hilbert spaces (with dimensions $d_j := \dim \mathcal{H}_j < \infty$), and $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_K$ their tensor product (of dimension $D := d_1 \cdots d_K$). Let furthermore Δ be a Hermitian operator on \mathcal{H} . The first measurements we shall be interested in, are tensor products of t -designs: $M = M^{(1)} \otimes \cdots \otimes M^{(K)}$, where each $M^{(j)}$ is a t -design POVM. In other words, the individual elements of the partition of unity are all possible tensor products $M_{x_1}^{(1)} \otimes \cdots \otimes M_{x_K}^{(K)}$. The following observation makes it possible to use probabilistic techniques to analyse the norm associated to a single measurement, paving the way to an analysis of t -design measurements.

Observation 2 For a rank-one POVM $M = (M_x)_{x \in \mathcal{X}}$ on a d -dimensional Hilbert space \mathcal{H} , let $p_x = \frac{1}{d} \text{Tr } M_x$ and $P_x = \frac{M_x}{\text{Tr } M_x}$. Then, introducing a random index X with $\Pr\{X = x\} = p_x$, P_X is a random rank-one projector with expectation $\mathbb{E}P_X = \sum_{x \in \mathcal{X}} p_x P_x = \frac{1}{d} \mathbb{1}$. Furthermore,

$$\|\Delta\|_M = d \mathbb{E}|S|,$$

for the real random variable $S = \text{Tr } \Delta P_X$. Indeed,

$$d \mathbb{E}|S| = d \sum_{x \in \mathcal{X}} p_x |\text{Tr } \Delta P_x| = \sum_{x \in \mathcal{X}} |\text{Tr } \Delta M_x| = \|\Delta\|_M.$$

Beyond these t -design tensor products, we are going to consider the class of all POVMs implementable by a protocol of local operations and classical communication (LOCC) which includes the above; the class SEP consisting of all POVMs $M = (M_x)_{x \in \mathcal{X}}$ with fully separable operators $M_x \geq 0$, which in turn contains LOCC; and finally the even larger class PPT that is defined by $M_x^{\Gamma_I} \geq 0$ for all $x \in \mathcal{X}$ and $I \subset [K]$, where Γ_I is the partial transpose on all parties I . By the definition of these classes, it is enough to consider two-outcome POVMs $(M, \mathbb{1} - M)$ that can be implemented by LOCC, or such that both M and $\mathbb{1} - M$ are separable, or PPT with respect to all bipartite cuts, respectively. See [12] for a more detailed discussion of these classes.

Definition 3 For any operator Δ (we only consider Hermitian ones in the following) on $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_K$, let

$$\|\Delta\|_{2(K)} := \sqrt{\sum_{I \subset [K]} \text{Tr} |\text{Tr}_I \Delta|^2},$$

where Tr_I denotes the partial trace over all parties I .

Note that for $K = 1$, this is “almost” the non-commutative ℓ_2 -norm: $\|\Delta\|_{2(1)} = \sqrt{|\text{Tr } \Delta|^2 + \text{Tr } |\Delta|^2}$, reducing to the latter (aka Hilbert-Schmidt norm), $\|\Delta\|_2 = \sqrt{\text{Tr } \Delta \Delta^\dagger}$, on traceless operators.

In [12] such measurements and the above classes LOCC, SEP and PPT were investigated in the case of $K = 1$ and $K = 2$ parties. Measurement norm and 2-norm were first directly related in [1], with an application in quantum algorithms, while Harrow *et al.* [8] were the first to realise that for a 4-design POVM M , the measurement norm and ℓ_2 -norm are indeed equivalent, although only for traceless operators Δ on a single system:

$$\|\Delta\|_2 \geq \|\Delta\|_M \geq \frac{1}{3} \|\Delta\|_2,$$

The extension to two parties in [12],

$$\|\Delta\|_M \geq \frac{1}{\sqrt{153}} \|\Delta\|_2,$$

for a tensor product of two 4-design POVMs and still assuming $\text{Tr } \Delta = 0$, subsequently found applications in entanglement theory [4], suggesting that our results for larger K might be useful, too.

III. COMPARISON WITH 2-NORMS

Our first two theorems show that the norms related to 2- and 4-designs are closely related to the norm $\|\cdot\|_{2(K)}$.

Theorem 4 *If M is a tensor product of K 2-design POVMs, then*

$$\|\Delta\|_M \leq \sqrt{\prod_{j=1}^K \frac{d_j}{d_j + 1}} \|\Delta\|_{2(K)} \leq \|\Delta\|_{2(K)}.$$

Proof Starting from observation 2, with the random variable S that takes the value $\text{Tr } \Delta P_{\underline{x}}$ with probability $p_{\underline{x}} = p_{x_1} \cdots p_{x_K}$, we have by the convexity of the square function,

$$\|\Delta\|_M = D \mathbb{E}|S| \leq D \sqrt{\mathbb{E}S^2}.$$

Furthermore, using the definition of 2-design,

$$\begin{aligned} \mathbb{E}S^2 &= \sum_{\underline{x}} p_{\underline{x}} (\text{Tr } \Delta P_{\underline{x}})^2 \\ &= \sum_{\underline{x}} p_{\underline{x}} \text{Tr}(\Delta \otimes \Delta)(P_{\underline{x}} \otimes P_{\underline{x}}) \\ &= \text{Tr} \left(\Delta^{\otimes 2} \bigotimes_{j=1}^K \frac{\mathbb{1} + F}{d_j(d_j + 1)} \right) \\ &= \prod_{j=1}^K \frac{1}{d_j(d_j + 1)} \sum_{I \subset [K]} \text{Tr}(\text{Tr}_I \Delta)^2, \end{aligned}$$

with the notation $F := U_{(12)}$, and using the fact that $\text{Tr}(A \otimes B)F = \text{Tr } AB$. Inserting this into the above inequality concludes the proof. \square

Theorem 5 *If M is a tensor product of K 4-design POVMs, then*

$$\sqrt{\frac{1}{18}} \|\Delta\|_{2(K)} \leq \|\Delta\|_M \leq \|\Delta\|_{2(K)}.$$

Proof Again we start with observation 2, with the random variable S that takes the value $\text{Tr } \Delta P_{\underline{x}}$ with probability $p_{\underline{x}} = p_{x_1} \cdots p_{x_K}$.

The upper bound is contained in theorem 4, as a t -design is automatically a $(t-1)$ -design. For the lower bound, we follow the strategy of Ambainis and Emerson [1], using this inequality of Berger's [3] (by the way a special case of Hölder's inequality):

$$\mathbb{E}|S| \geq \sqrt{\frac{(\mathbb{E}S^2)^3}{\mathbb{E}S^4}}.$$

In the proof of theorem 4 we have already calculated

$$\mathbb{E}S^2 = \prod_{j=1}^K \frac{1}{d_j(d_j+1)} \sum_{I \subset [K]} \text{Tr}(\text{Tr}_I \Delta)^2.$$

Using the property of 4-design, we similarly get

$$\mathbb{E}S^4 = \prod_{j=1}^K \frac{1}{d_j(d_j+1)(d_j+2)(d_j+3)} \text{Tr} \left(\Delta^{\otimes 4} \left(\sum_{\underline{\pi} \in \mathfrak{S}_4^K} U_{\underline{\pi}} \right) \right),$$

with the notation $U_{\underline{\pi}} := \bigotimes_{j=1}^K U_{\pi_j}$ for $\underline{\pi} = (\pi_1, \dots, \pi_K)$.

Thus it suffices to show

$$\text{Tr} \left(\Delta^{\otimes 4} \left(\sum_{\underline{\pi} \in \mathfrak{S}_4^K} U_{\underline{\pi}} \right) \right) \leq 18^K \left[\sum_{I \subset [K]} \text{Tr}(\text{Tr}_I \Delta)^2 \right]^2,$$

which is precisely proposition 11 in appendix A, and we are done. \square

Alternative proof of a weaker version of theorem 5. Here is a way of demonstrating the slightly worse bound

$$\text{Tr} \left(\Delta^{\otimes 4} \left(\sum_{\underline{\pi} \in \mathfrak{S}_4^K} U_{\underline{\pi}} \right) \right) \leq 24^K \left[\sum_{I \subset [K]} \text{Tr}(\text{Tr}_I \Delta)^2 \right]^2,$$

which has the advantage of being conceptually simple, and showing some of the tricks used in the proof of proposition 11. For this it is enough to show that, for every K -tuple $\underline{\pi} \in \mathfrak{S}_4^K$:

$$t(\underline{\pi}) := |\text{Tr } \Delta^{\otimes 4} U_{\underline{\pi}}| \leq \max_{I \subset [K]} \left[\text{Tr}(\text{Tr}_I \Delta)^2 \right]^2. \quad (1)$$

The basic idea is to use Cauchy-Schwarz inequality repeatedly, as in [12]: For arbitrary (compatible) operators X and Y ,

$$|\text{Tr } XY^\dagger| \leq \sqrt{(\text{Tr } XX^\dagger)(\text{Tr } YY^\dagger)}.$$

Concretely, for Hermitian operators M_1, M_2, M_3, M_4 on a Hilbert space \mathcal{K} and a permutation $\sigma \in \mathfrak{S}_4$ with corresponding unitary U_σ on $\mathcal{K}^{\otimes 4}$, we may write

$$\text{Tr } U_\sigma(M_1 \otimes M_2 \otimes M_3 \otimes M_4) = \text{Tr } XY^\dagger,$$

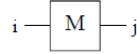
with operators X and Y and permutations $\sigma^L, \sigma^R \in \mathfrak{S}_4$ such that

$$\text{Tr } XX^\dagger = \text{Tr } U_{\sigma^L}(M_1 \otimes M_2 \otimes M_2 \otimes M_1),$$

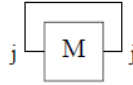
$$\text{Tr } YY^\dagger = \text{Tr } U_{\sigma^R}(M_4 \otimes M_3 \otimes M_3 \otimes M_4).$$

What we gain by doing so is that σ^L and σ^R are not arbitrary elements of \mathfrak{S}_4 ; rather, they necessarily belong to the subset $\mathfrak{A} := \{\text{id}, (14), (23), (1234), (1432), (12)(34), (14)(23)\}$, which is stable under the exchange $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$, i.e. under the conjugation by $(14)(23)$. In order to easily see into which pair $(\sigma^L, \sigma^R) \in \mathfrak{A} \times \mathfrak{A}$ each permutation $\sigma \in \mathfrak{S}_4$ splits, we can make use of Penrose's tensor diagrams [14], which we briefly explain here.

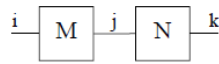
For any Hermitian M on \mathcal{K} and unit vectors $|i\rangle, |j\rangle \in \mathcal{K}$ we represent the matrix element $\langle i|M|j\rangle$ by the following diagram with terminals:



Summing matrix elements over an orthonormal basis of \mathcal{K} is represented by joining the corresponding terminals. So, for instance, $\text{Tr } M = \sum_j \langle j|M|j\rangle$ is represented by



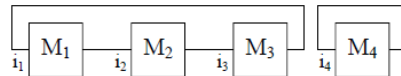
And in the same way for matrix multiplication, $\langle i|M|j\rangle\langle j|N|k\rangle$ is represented by



The expressions we looked at above are, for Hermitian M_1, M_2, M_3, M_4 on \mathcal{K} and $\sigma \in \mathfrak{S}_4$:

$$\text{Tr } U_\sigma(M_1 \otimes M_2 \otimes M_3 \otimes M_4) = \sum_{i_1, i_2, i_3, i_4} \langle i_1|M_1|i_{\sigma(1)}\rangle \langle i_2|M_2|i_{\sigma(2)}\rangle \langle i_3|M_3|i_{\sigma(3)}\rangle \langle i_4|M_4|i_{\sigma(4)}\rangle.$$

For instance, $\text{Tr } U_{(123)}(M_1 \otimes M_2 \otimes M_3 \otimes M_4)$ is represented by



In this case, the splitting procedure and use of Cauchy-Schwarz described above can be schematically written as

$$\left| \begin{array}{c} \text{Diagram 1} \\ \text{Diagram 2} \end{array} \right| \leq \sqrt{\begin{array}{c} \text{Diagram 3} \\ \text{Diagram 4} \end{array}}$$

Diagram 1: A sequence of boxes M1, M2, M3, M4. M1 and M2 are connected by a horizontal line labeled X. M3 and M4 are connected by a horizontal line labeled Y†. A vertical dashed line separates M2 and M3.

Diagram 2: A sequence of boxes M1, M2, M2, M1. M1 and M2 are connected by a horizontal line labeled X. M2 and M1 are connected by a horizontal line labeled X†.

Diagram 3: A sequence of boxes M4, M3, M3, M4. M4 and M3 are connected by a horizontal line labeled Y. M3 and M4 are connected by a horizontal line labeled Y†.

Diagram 4: A sequence of boxes M1, M2, M2, M1. M1 and M2 are connected by a horizontal line labeled X. M2 and M1 are connected by a horizontal line labeled X†.

Conj. class	σ	σ^L	σ^R
1111	id	id	id
211	(12)	(12)(34)	id
	(13)	(14)	(23)
	(14)	(14)	(14)
	(23)	(23)	(23)
	(24)	(23)	(14)
	(34)	id	(12)(34)
22	(12)(34)	(12)(34)	(12)(34)
	(13)(24)	(14)(23)	(14)(23)
	(14)(23)	(14)(23)	(14)(23)
31	(123)	(1234)	(23)
	(132)	(1432)	(23)
	(124)	(1234)	(14)
	(142)	(1432)	(14)
	(134)	(14)	(1234)
	(143)	(14)	(1432)
	(234)	(23)	(1234)
	(243)	(23)	(1432)
4	(1234)	(1234)	(1234)
	(1243)	(1234)	(1432)
	(1324)	(14)(23)	(14)(23)
	(1342)	(1432)	(1234)
	(1432)	(1432)	(1432)
	(1423)	(14)(23)	(14)(23)

FIG. 1. Table of the splitting map $\text{Split} : \mathfrak{S}_4 \longrightarrow \mathfrak{A} \times \mathfrak{A}$, $\text{Split}(\sigma) = (\sigma^L, \sigma^R)$, grouped according to conjugacy classes of σ .

which means that for $\sigma = (123)$, we have $\sigma^L = (1234)$ and $\sigma^R = (23)$.

The resulting splitting map $\text{Split} : \mathfrak{S}_4 \ni \sigma \mapsto (\sigma^L, \sigma^R) \in \mathfrak{A} \times \mathfrak{A}$ for each $\sigma \in \mathfrak{S}_4$ can then easily be constructed and looked up in the table of Fig. 1.

Trivially extending the above reasoning to K -tuples $\underline{\pi} = (\pi_1, \dots, \pi_K) \in \mathfrak{S}_4^K$ of permutations, we apply the splitting map to all the π_i ($1 \leq i \leq K$), and use the Cauchy-Schwarz as well as geometric-arithmetic mean inequality:

$$\begin{aligned}
 t(\underline{\pi}) = |\text{Tr } \Delta^{\otimes 4} U_{\underline{\pi}}| &\leq \sqrt{(\text{Tr } \Delta^{\otimes 4} U_{\underline{\pi}^L}) (\text{Tr } \Delta^{\otimes 4} U_{\underline{\pi}^R})} = \sqrt{t(\underline{\pi}^L) t(\underline{\pi}^R)} \\
 &\leq \frac{1}{2} t(\underline{\pi}^L) + \frac{1}{2} t(\underline{\pi}^R).
 \end{aligned} \tag{2}$$

The other observation we use is that $t(\underline{\pi})$ is invariant under conjugation by elements from the diagonal subgroup $\mathfrak{G} := \{(\sigma, \dots, \sigma), \sigma \in \mathfrak{S}_4\}$ of \mathfrak{S}_4^K , because $\Delta^{\otimes 4}$ is invariant under conjugation by elements of the form $(U_\sigma)^{\otimes K}$. Now, notice that the subset $\mathfrak{A}_0 = \{\text{id}, (12)(34), (14)(23)\}$ of \mathfrak{A} is such that \mathfrak{A}_0^K is stable under conjugation by any element of \mathfrak{G} followed by splitting. And what is more, any given $\underline{\pi} \in \mathfrak{S}_4^K$ can be transformed into a family of elements of \mathfrak{A}_0 by repeatedly conjugating by elements of \mathfrak{G} and splitting.

Thus, using eq. (2) and conjugation invariance repeatedly, we eventually get for all $\underline{\pi} \in \mathfrak{S}_4^K$ the upper bound

$$t(\underline{\pi}) \leq \sum_{\alpha} p_{\alpha} t(\underline{\pi}^{(\alpha)}),$$

with certain $p_{\alpha} \geq 0$ summing to 1, and $\underline{\pi}^{(\alpha)}$ belonging to \mathfrak{A}_0^K . As we have not attempted to control the coefficients p_{α} , we record as a useful intermediate bound for all $\underline{\pi} \in \mathfrak{S}_4^K$,

$$t(\underline{\pi}) \leq \max_{\underline{\sigma} \in \mathcal{A}_0^K} t(\underline{\sigma}). \quad (3)$$

As a matter of fact, we know already how to upper bound the traces on the right hand side of eq. (3). Indeed, a generic $\underline{\sigma} \in \mathcal{A}_0^K$ is given by disjoint subsets $I, J \subset [K]$, such that:

$$\sigma_j = \begin{cases} \text{id} & j \in I, \\ (12)(34) & j \in J, \\ (14)(23) & j \in J' := [K] \setminus (I \cup J). \end{cases}$$

Hence,

$$t(\underline{\sigma}) = \text{Tr } \Delta^{\otimes 4} U_{\underline{\sigma}} = \text{Tr} \left((\text{Tr}_I \Delta)^{\otimes 4} (U_{(12)(34)}^{\otimes |J|} \otimes U_{(14)(23)}^{\otimes |J'|}) \right),$$

which really is a bipartite term (i.e., $K = 2$) as treated in [12, Proof of Lemma 26, case “(2,2):(2,2)”]: there it was shown to be $\leq [\text{Tr}(\text{Tr}_I \Delta)^2]^2$.

Combining this bound with eq. (3), we obtain eq. (1), and we are done. \square

Theorem 5 extends the results of [1, 12] to $K > 2$; however we may wonder how good the lower bound really is, and in particular if the dependence on K is “real”. The following result shows that, indeed, the constant relating $\|\Delta\|_M$ has to decrease as a power of K . For this it is enough to analyse a specific tensor product of local 4-design POVMs, and we choose $U_{\mathcal{H}} := U_{\mathcal{H}_1} \otimes \cdots \otimes U_{\mathcal{H}_K}$, the tensor product of the K uniform (unitary invariant) POVMs on sub-systems \mathcal{H}_j ($j = 1, \dots, K$). This is an interesting measurement since each of the $U_{\mathcal{H}_j}$ is an ∞ -design, in particular a 4-design, and we can exploit the symmetry to make calculations feasible. Whereas theorem 5 gives us

$$\sqrt{\frac{1}{18}}^K \|\Delta\|_2 \leq \sqrt{\frac{1}{18}}^K \|\Delta\|_{2(K)} \leq \|\Delta\|_{U_{\mathcal{H}}} \leq \|\Delta\|_{2(K)},$$

we have the following:

Proposition 6 *There exists a Hermitian $\Delta \neq 0$ on \mathcal{H} such that*

$$\|\Delta\|_{U_{\mathcal{H}}} = \sqrt{\frac{1}{2}}^K \|\Delta\|_{2(K)} = \sqrt{\frac{1}{2}}^K \|\Delta\|_2.$$

Proof Define $\Delta_j = \frac{1}{2}|0\rangle\langle 0| - \frac{1}{2}|1\rangle\langle 1|$ with orthogonal unit vectors $|0\rangle, |1\rangle \in \mathcal{H}_j$, $j = 1, \dots, K$. Let $\Delta = \bigotimes_{j=1}^K \Delta_j$. Clearly, $\text{Tr } \Delta_j = 0$ and $\|\Delta_j\|_2 = \sqrt{\frac{1}{2}}$ for all j , while from [12, Theorem 10] we know that $\|\Delta_j\|_{U_{\mathcal{H}_j}} = \frac{1}{2}$.

Hence, $\|\Delta\|_{2(K)} = \|\Delta\|_2 = 2^{-\frac{K}{2}}$; on the other hand, exploiting the tensor product structure of both state and measurement, $\|\Delta\|_{U_{\mathcal{H}}} = 2^{-K}$. \square

We shall now move on to investigating the properties of the measurement norms associated with not one but a whole class of locally restricted measurements.

Theorem 7 For any number $K \geq 2$ of parties and any local dimensions d_j ($1 \leq j \leq K$),

$$\|\Delta\|_{\text{SEP}} \geq 2\sqrt{\frac{1}{2}}^K \|\Delta\|_2.$$

Furthermore,

$$\|\Delta\|_{\text{PPT}} \geq \|\Delta\|_2.$$

Proof The first inequality above was already shown in [12], but we repeat the proof since it is very simple: It uses a result of Barnum and Gurvits [2], that for any Hermitian X on a K -partite Hilbert space, if $\|X - \mathbb{1}\|_2 \leq 2^{1-\frac{K}{2}}$, then X is separable.

So, $\|2M - \mathbb{1}\|_2 \leq 2^{1-\frac{K}{2}}$ implies that both $\mathbb{1} + (2M - \mathbb{1}) = 2M$ and $\mathbb{1} + (\mathbb{1} - 2M) = 2(\mathbb{1} - M)$ are separable, i.e. M and $\mathbb{1} - M$ are separable operators. Thus, for our Hermitian Δ on \mathcal{H} :

$$\begin{aligned} \|\Delta\|_{\text{SEP}} &= \max_{(M, \mathbb{1}-M) \in \text{SEP}} |\text{Tr } \Delta(2M - \mathbb{1})| \\ &\geq \max_{\|A\|_2 \leq 2^{1-\frac{K}{2}}} |\text{Tr } \Delta A| \\ &= 2^{1-\frac{K}{2}} \|\Delta\|_2, \end{aligned}$$

where the last equality is by self-duality of the ℓ_2 -norm.

To show the second inequality, notice that $(M, \mathbb{1} - M)$ being a two-outcome PPT POVM is a consequence of M and $\mathbb{1} - M$ being separable for any bipartition of the K parties.

Thus, we can use once more the Barnum-Gurvits result [2]: If $\|2M - \mathbb{1}\|_2 \leq 1$, then M and $\mathbb{1} - M$ are both separable with respect to any bipartition, hence PPT with respect to any bipartition. The claim follows now as in the first part. \square

IV. COMPARISON WITH TRACE NORM AND DATA HIDING

All measurement norms are trivially upper bounded by the trace norm $\|\cdot\|_1$. In the other direction, the standard $\|\Delta\|_1 \leq \sqrt{D}\|\Delta\|_2$ for operators Δ on a D -dimensional Hilbert space, allows us to turn the ℓ_2 -norm estimates from the previous section into lower bounds on $\|\Delta\|_M$, which in turn provides a lower bound on $\|\Delta\|_{\text{LOCC}} \leq \|\Delta\|_{\text{SEP}} \leq \|\Delta\|_{\text{PPT}}$:

$$\|\Delta\|_M \geq \sqrt{\frac{1}{18}}^K \frac{1}{\sqrt{D}} \|\Delta\|_1, \quad (4)$$

for any tensor product of 4-design POVMs, M . These are non-trivial because by now it is a classic result in quantum information that quantum states allow for *data hiding* [16]: namely, on large composite systems there exist states with orthogonal supports (hence perfect distinguishability by a suitable measurement) that are nevertheless barely distinguishable by LOCC.

More of this below, but let us start with some simple observations: That both the occurrence of the inverse square root of D , and the exponential dependence of the lower bound on K are not artifacts, is shown by the following example.

Proposition 8 Consider the measurement $U_{\mathcal{H}} := U_{\mathcal{H}_1} \otimes \cdots \otimes U_{\mathcal{H}_K}$, the tensor product of the K uniform (unitary invariant) POVMs on sub-systems \mathcal{H}_j with dimensions d_j , $j = 1, \dots, K$. There exists a Hermitian $\Delta \neq 0$ such that

$$\|\Delta\|_{U_{\mathcal{H}}} \leq \left(\sqrt{\frac{2}{\pi}} + o(1) \right)^K \frac{1}{\sqrt{D}} \|\Delta\|_1,$$

where $o(1)$ is arbitrarily small for sufficiently large $d_{\min} = \min\{d_1, \dots, d_K\}$.

Proof Without loss of generality, all d_j are even. Pick any projector P_j of rank $\frac{d_j}{2}$ in \mathcal{H}_j , $Q_j := \mathbb{1} - P_j$, and let $\Delta_j := \frac{1}{d_j}P_j - \frac{1}{d_j}Q_j$, so that $\|\Delta_j\|_1 = 1$.

Our candidate is $\Delta = \bigotimes_{j=1}^K \Delta_j$, which also has trace norm 1. On the other hand, by [12, Theorem 10] we have

$$\|\Delta_j\|_{U_{\mathcal{H}_j}} \leq \left(\sqrt{\frac{2}{\pi}} + o(1) \right) \frac{1}{\sqrt{d_j}},$$

and since both Δ and the measurement share the tensor product structure, we obtain the claim by multiplying together these inequalities. \square

That the factor of $\frac{1}{\sqrt{D}}$ does not go away when we go to the class of all LOCC, and indeed all PPT measurements, is contained in the two following theorems.

Theorem 9 *When all the local dimensions are d , hence $D = d^K$, there exists a traceless Hermitian $\Delta \neq 0$ with*

$$\|\Delta\|_{\text{PPT}} \leq \frac{2}{d^{\lfloor K/2 \rfloor} - 1} \|\Delta\|_1.$$

In other words, one can find two states ρ_0 and ρ_1 with orthogonal supports (i.e., $\frac{1}{2}\rho_0 - \frac{1}{2}\rho_1$ has trace norm 1), such that

$$\left\| \frac{1}{2}\rho_0 - \frac{1}{2}\rho_1 \right\|_{\text{PPT}} \leq \frac{2\sqrt{d}^\kappa}{\sqrt{D} - \sqrt{d}^\kappa} \leq \frac{3\sqrt{d}^\kappa}{\sqrt{D}},$$

where $\kappa = K \bmod 2$ is the parity of K . Hence these two states are data hiding in the sense of [16]: ρ_i encodes a state between K parties, but as long as those are restricted to LOCC measurements (or more generally PPT measurements), they have only a very slim chance of guessing this state. Indeed, the probability of discriminating correctly ρ_0 from ρ_1 decreases as the inverse square root of the total dimension D , and eq. (4) shows that this order of magnitude is essentially optimal, apart from a K -dependent constant.

Proof For all Hermitian Δ ,

$$\begin{aligned} \|\Delta\|_{\text{PPT}} &= \max_{(\frac{1}{2}(\mathbb{1}+A), \frac{1}{2}(\mathbb{1}-A)) \in \text{PPT}} |\text{Tr } A\Delta| \\ &= \max_{\forall I \subset [K], -\mathbb{1} \leq A^{\Gamma_I} \leq \mathbb{1}} |\text{Tr } A\Delta|. \end{aligned}$$

Yet, if A is such that for $I \subset [K]$, $-\mathbb{1} \leq A^{\Gamma_I} \leq \mathbb{1}$, then necessarily

$$|\text{Tr } A\Delta| = |\text{Tr } A^{\Gamma_I} \Delta^{\Gamma_I}| \leq \|A^{\Gamma_I}\|_\infty \|\Delta^{\Gamma_I}\|_1 \leq \|\Delta^{\Gamma_I}\|_1.$$

Among the operators for which we know how to evaluate the trace norm of any of their partial transposes are the permutation operators U_π , $\pi \in \mathfrak{S}_K$. Indeed, for all $I := \{1, \dots, p\} \subset [K]$ we have

$$U_\pi^{\Gamma_I} = \sum_{j_1, \dots, j_K} |j_{\pi(1)}, \dots, j_{\pi(p)}, j_{p+1}, \dots, j_K\rangle \langle j_1, \dots, j_p, j_{\pi(p+1)}, \dots, j_{\pi(K)}|.$$

Hence, letting $f(I, \pi) := |\{i \in I, \pi(i) \notin I\}|$, we get: $\|U_\pi^{\Gamma_I}\|_1 = d^{K-f(I, \pi)}$.

Choosing as permutation π the product of $\lfloor K/2 \rfloor$ disjoint transpositions, $\pi := (1, \lfloor K/2 \rfloor + 1) \dots (\lfloor K/2 \rfloor, 2\lfloor K/2 \rfloor)$ (that decomposes therefore into $\lceil K/2 \rceil$ disjoint cycles), let us now consider the following traceless Hermitian Δ :

$$\begin{aligned} \Delta &:= \frac{1}{d^K + d^{\lceil K/2 \rceil}}(\mathbb{1} + U_\pi) - \frac{1}{d^K - d^{\lceil K/2 \rceil}}(\mathbb{1} - U_\pi) \\ &= \frac{2}{d^{\lceil K/2 \rceil}(d^{2\lceil K/2 \rceil} - 1)} \left(d^{\lfloor K/2 \rfloor} U_\pi - \mathbb{1} \right). \end{aligned}$$

Note that Δ is the difference of the two orthogonal density operators $\rho_0 := \frac{1}{d^K + d^{\lceil K/2 \rceil}}(\mathbb{1} + U_\pi)$ and $\rho_1 := \frac{1}{d^K - d^{\lceil K/2 \rceil}}(\mathbb{1} - U_\pi)$, hence $\|\Delta\|_1 = 2$.

Furthermore, $I := \{1, \dots, \lfloor K/2 \rfloor\} \subset [K]$ is such that $f(I, \pi) = \lfloor K/2 \rfloor$, so $\|U_\pi^{\Gamma_I}\|_1 = d^{K-\lfloor K/2 \rfloor}$, and hence, after a straightforward calculation:

$$\|\Delta^{\Gamma_I}\|_1 \leq \frac{2}{d^{\lceil K/2 \rceil}(d^{2\lceil K/2 \rceil} - 1)} \left(d^{\lfloor K/2 \rfloor} \|U_\pi^{\Gamma_I}\|_1 + \|\mathbb{1}^{\Gamma_I}\|_1 \right) \leq \frac{2}{d^{\lfloor K/2 \rfloor} - 1} \|\Delta\|_1.$$

Thus, $\|\Delta\|_{\text{PPT}} \leq \|\Delta^{\Gamma_I}\|_1 \leq \frac{2}{d^{\lfloor K/2 \rfloor} - 1} \|\Delta\|_1$, which is what we wanted to prove. \square

Theorem 9 and eq. (4) show that – at least for even K – the best performance for K -party data hiding is indeed a bias inversely proportional to the square root of the dimension, with a constant factor only depending on K . Here is another construction that works also for odd number K of parties, with possibly unequal local dimensions.

Theorem 10 *When all the local dimensions d_j ($1 \leq j \leq K$) are such that there exists a $I \subset [K]$ such that $\mathcal{A} := \bigotimes_{j \in I} \mathcal{H}_j$ and $\mathcal{B} := \bigotimes_{j \in [K] \setminus I} \mathcal{H}_j$ satisfy $\dim \mathcal{A} = \dim \mathcal{B} = \sqrt{D}$, then there exists a traceless Hermitian $\Delta \neq 0$ with*

$$\|\Delta\|_{\text{PPT}} \leq \frac{2}{\sqrt{D} + 1} \|\Delta\|_1.$$

Proof Denoting by F the swap operator between the Hilbert spaces \mathcal{A} and \mathcal{B} , we let σ and α be the normalised projectors onto the symmetric and antisymmetric subspaces of $\mathbb{C}^{\sqrt{D}} \otimes \mathbb{C}^{\sqrt{D}}$, respectively: $\sigma := \frac{1}{\sqrt{D}(\sqrt{D}+1)}(\mathbb{1} + F)$ and $\alpha := \frac{1}{\sqrt{D}(\sqrt{D}-1)}(\mathbb{1} - F)$. We then consider the traceless Hermitian $\Delta := \sigma - \alpha$.

Now, if a POVM is PPT across all possible bipartitions of \mathcal{H} , it is in particular PPT across the bipartition $\mathcal{A} : \mathcal{B}$. As a consequence,

$$\|\Delta\|_{\text{PPT}} \leq \|\Delta\|_{\text{PPT}(\mathcal{A}:\mathcal{B})} = \frac{2}{\sqrt{D} + 1} \|\Delta\|_1,$$

where the last equality is the original quantum data hiding result, as shown in [5, 16]. \square

V. CONCLUSION

We have solved an open problem from [12], showing that for any number K of parties, the measurement norm on Hermitian operators defined by local 4-designs is equivalent to a certain relative of the Hilbert-Schmidt norm. The equivalence is in terms of constants of domination which depend only on the number of parties, not on the local dimensions.

$$\begin{array}{ccc}
\frac{1}{\sqrt{D}} \|\Delta\|_1 \leq & \|\Delta\|_2 \leq & \|\Delta\|_{\text{PPT}} \leq \|\Delta\|_1 \\
& & \forall I \\
2\sqrt{\frac{1}{2}} \frac{1}{\sqrt{D}} \|\Delta\|_1 \leq 2\sqrt{\frac{1}{2}} \|\Delta\|_2 \leq & & \|\Delta\|_{\text{SEP}} \\
& & \forall I \\
& & \|\Delta\|_{\text{LOCC}} \\
& & \forall I \\
\sqrt{\frac{1}{18}} \frac{1}{\sqrt{D}} \|\Delta\|_1 \leq \sqrt{\frac{1}{18}} \|\Delta\|_2 \leq \sqrt{\frac{1}{18}} \|\Delta\|_{2(K)} \leq \|\Delta\|_{M_4^{(K)}} \leq \|\Delta\|_{2(K)}
\end{array}$$

FIG. 2. A schematic summary of the new and previously known relations; $M_4^{(K)}$ denotes a generic tensor product of 4-design POVMs.

Note that our constants appear worse compared to the known inequalities for $K = 1$ and $K = 2$: In the former case, [1] gives $\frac{1}{3}$ whereas we get $\sqrt{\frac{1}{18}}$; in the latter, [12] gives $\frac{1}{\sqrt{153}}$ whereas we get $\frac{1}{18}$. While the gap is small, it may to some degree be explained by the fact that in both these cited papers the assumption $\text{Tr } \Delta = 0$ was made, and exploited to simplify the fourth moment even more. We believe that there is merit in transcending this restriction, as not in all applications it can be justified. In any case, we leave it as an open problem to find the optimal constants of domination with respect to the $\|\cdot\|_{2(K)}$ norm.

Via the non-commutative ℓ_2 -norm we then obtained performance comparisons with the trace norm, revealing at most a factor of the order of the inverse square root of the dimension of the total Hilbert space between the measurement norm and the trace norm. Since the measurement is a particular LOCC strategy, we get lower bounds on the distinguishing power of LOCC measurements. The bounds can be shown to be optimal in their dimensional dependence, as we exhibited two constructions of data hiding states which attain these bounds up to K -dependent factors. Here, one remaining question is whether for odd number K of parties, all of which have equal dimension, the additional factor of square root of the local dimension can be removed in theorem 9. On a related note, with respect to theorem 10, does there exist a universal constant $C > 0$ such that for all sufficiently large D one can find Hermitian $\Delta \neq 0$ with $\|\Delta\|_{\text{PPT}} \leq \frac{C}{\sqrt{D}} \|\Delta\|_1$, irrespective of the local dimensions?

Even more interesting would be to quantify the performance of LOCC, or at least fully separable (SEP), measurements: Indeed, notice that in theorems 9 and 10, we have only exploited *bi*-separability, and comparing with theorem 7 we see that there remains only a factor of at most 2 to be gained as long as one is restricted to this weaker constraint. Is it possible to significantly improve this factor when judging the performance of SEP or LOCC measurements? In particular, do there exist constants $C > 0$ and $\alpha < 1$ such that for all K and all sufficiently large total dimensions D there is a Hermitian $\Delta \neq 0$ with

$$\|\Delta\|_{\text{LOCC}} \leq C \frac{\alpha^K}{\sqrt{D}} \|\Delta\|_1, \text{ or even } \|\Delta\|_{\text{SEP}} \leq C \frac{\alpha^K}{\sqrt{D}} \|\Delta\|_1 ?$$

ACKNOWLEDGMENTS

We thank Fernando Brandão, Aram Harrow and Ashley Montanaro for spurring our interest in generalising the 4-design results of [1, 12] to multiple parties. Their observations directed our attention towards 2-norms as the correct framework for understanding state discrimination with a fixed measurement.

This work was done as part of CL's research placement at the University of Bristol. AW is supported by the European Commission (STREP "QCS" and Integrated Project "QESSENCE"), the ERC (Advanced Grant "IRQUAT"), a Royal Society Wolfson Merit Award and a Philip Leverhulme Prize. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

-
- [1] A. Ambainis, J. Emerson, "Quantum t-designs: t-wise independence in the quantum world", Proc. 22nd IEEE Conf. Computational Complexity (CCC07), pp. 129140, IEEE, Piscataway, NJ, 2007; arXiv:quant-ph/0701126v2.
 - [2] H. N. Barnum, L. Gurvits, "Separable balls around the maximally mixed multipartite quantum states", Phys. Rev. A 68:042312 (2003); arXiv:quant-ph/0302102v1.
 - [3] B. Berger, "The fourth moment method", SIAM J. Comput. 24(6):1188-1207 (1997).
 - [4] F. G. S. L. Brandão, M. Christandl, J. T. Yard, "Faithful Squashed Entanglement", Commun. Math. Phys. 306:805-830 (2011); arXiv[quant-ph]:1010.1750v5.
 - [5] T. Eggeling, R. F. Werner, "Hiding classical data in multi-partite quantum states"; Phys. Rev. Lett. 89:097905 (2003); arXiv:quant-ph/0203004v2.
 - [6] W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley Series in Probability and Mathematical Statistics, New York, 1966.
 - [7] D. Gross, K. Audenaert, J. Eisert, "Evenly distributed unitaries: on the structure of unitary designs", J. Math. Phys. 48:052104 (2007); arXiv:quant-ph/0611002v2.
 - [8] A. W. Harrow, A. Montanaro, A. J. Short, "Limitations on quantum dimensionality reduction", Proc. ICALP'11, LNCS 6755, pp.86-97, Springer Verlag, Berlin Heidelberg, 2011; arXiv[quant-ph]:1012.2262v2.
 - [9] C. W. Helstrom, *Quantum detection and estimation theory*, Academic Press, New York, 1976.
 - [10] A. S. Holevo, "Statistical decision theory for quantum systems", J. Multivariate Analysis 3:337-394 (1973).
 - [11] A. Klappenecker, M. Roetteler, "Mutually unbiased bases are complex projective 2-designs", Proc. ISIT 2005, pp. 1740-1744, IEEE, Piscataway, NJ, 2005; arXiv:quant-ph/0502031v2.
 - [12] W. Matthews, S. Wehner, A. Winter, "Distinguishability of quantum states under restricted families of measurements with an application to data hiding", Commun. Math. Phys. 291(3):813-843 (2009); arXiv[quant-ph]:0810.2327v2.
 - [13] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.
 - [14] R. Penrose, W. Rindler, *Spinors and Space-Time, Vol. 1: Two-spinor calculus and relativistic fields*, Cambridge University Press, Cambridge, 1986.
 - [15] J. M. Steele, *The Cauchy-Schwarz Master Class*, Cambridge University Press, Cambridge, 2004.
 - [16] B. M. Terhal, D. P. DiVincenzo, D. Leung, "Hiding Bits in Bell States", Phys. Rev. Lett. 86(25):5807-5810 (2001); arXiv:quant-ph/0011042v3. D. P. DiVincenzo, D. Leung, B. M. Terhal, "Quantum Data Hiding", IEEE Trans. Inf. Theory 48(3):580-599 (2002); arXiv:quant-ph/0103098v1.
 - [17] G. Zauner, *Quantum Designs -- Foundations of a Non-Commutative Theory of Designs* (in German). Ph.D. thesis, Universität Wien (1999).

Appendix A: Proof of theorem 5 – moment inequality

Here we show the missing ingredient, the following moment inequality:

Proposition 11 *For any Hermitian operator Δ on $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_K$,*

$$\mathrm{Tr} \left(\Delta^{\otimes 4} \left(\sum_{\pi \in \mathfrak{S}_4^K} U_\pi \right) \right) \leq 18^K \left[\sum_{I \subset [K]} \mathrm{Tr} (\mathrm{Tr}_I \Delta)^2 \right]^2.$$

In the proof of theorem 4 we could easily calculate

$$\mathrm{Tr} \left(\Delta^{\otimes 2} \left(\sum_{\pi \in \mathfrak{S}_2^K} U_\pi \right) \right) = \sum_{I \subset [K]} \mathrm{Tr} (\mathrm{Tr}_I \Delta)^2,$$

because \mathfrak{S}_2 only contains 2 elements. Now, \mathfrak{S}_4 contains 24 elements, so to upper bound

$$\mathrm{Tr} \left(\Delta^{\otimes 4} \left(\sum_{\pi \in \mathfrak{S}_4^K} U_\pi \right) \right)$$

we will find a way of restricting our attention to only a few of them without loss of generality. A strategy to do so has already been described when proving the weaker version of theorem 5 in section III.

What we have shown there is that, for all $\underline{\sigma} \in \mathfrak{S}_4^K$:

$$|\mathrm{Tr} (\Delta^{\otimes 4} U_{\underline{\sigma}})| \leq \frac{1}{2} \mathrm{Tr} (\Delta^{\otimes 4} U_{\underline{\sigma}^L}) + \frac{1}{2} \mathrm{Tr} (\Delta^{\otimes 4} U_{\underline{\sigma}^R}), \quad (\text{A1})$$

with $\underline{\sigma}^L, \underline{\sigma}^R \in \mathfrak{A}^K := \{\mathrm{id}, (14), (23), (1234), (1432), (12)(34), (14)(23)\}^K$ given by the splitting map detailed in the table in Fig. 1.

Consequently, in order to bound $|\mathrm{Tr} \Delta^{\otimes 4} U_{\underline{\sigma}}|$ for any $\underline{\sigma} \in \mathfrak{S}_4^K$, it will be sufficient to bound it for $\underline{\sigma} \in \mathfrak{A}^K$. Note that for the latter, the trace is automatically real and non-negative.

With this aim in view, let us first deal with the following auxiliary problem:

Let $\mathcal{H} = \mathcal{A} \otimes \cdots \otimes \mathcal{G}$ be a (finite dimensional) septempartite Hilbert space. For a generic operator X on \mathcal{H} and unit (typically: basis) vectors $|a\rangle, |a'\rangle \in \mathcal{A}, \dots, |g\rangle, |g'\rangle \in \mathcal{G}$, we denote by $X_{a,\dots,g}^{a',\dots,g'}$ the matrix element $\langle a | \cdots \langle g | X | a' \rangle \cdots | g' \rangle$.

Let $\underline{\sigma} = (\sigma_{\mathcal{A}}, \dots, \sigma_{\mathcal{G}}) \in \mathfrak{S}_4^7$ be a septuple of permutations. Now we have, with the $|a_q\rangle, \dots, |g_q\rangle$ ($1 \leq q \leq 4$) running over an orthonormal basis of $\mathcal{A}, \dots, \mathcal{G}$, respectively:

$$\mathrm{Tr} \Delta^{\otimes 4} (U_{\sigma_{\mathcal{A}}} \otimes \cdots \otimes U_{\sigma_{\mathcal{G}}}) = \sum_{\substack{a_1, \dots, g_1 \\ a_2, \dots, g_2 \\ a_3, \dots, g_3 \\ a_4, \dots, g_4}} \prod_{q=1}^4 \Delta_{a_q, \dots, g_q}^{a_{\sigma_{\mathcal{A}}(q)}, \dots, g_{\sigma_{\mathcal{G}}(q)}}.$$

In the particular case of all the seven permutations in \mathfrak{A} , $\sigma_{\mathcal{A}} = \text{id}$, $\sigma_{\mathcal{B}} = (14)$, $\sigma_{\mathcal{C}} = (23)$, $\sigma_{\mathcal{D}} = (1234)$, $\sigma_{\mathcal{E}} = (1432)$, $\sigma_{\mathcal{F}} = (12)(34)$ and $\sigma_{\mathcal{G}} = (14)(23)$, this becomes

$$\begin{aligned} \text{Tr } \Delta^{\otimes 4}(U_{\sigma_{\mathcal{A}}} \otimes \cdots \otimes U_{\sigma_{\mathcal{G}}}) &= \sum_{\substack{a_1, \dots, g_1 \\ a_2, \dots, g_2 \\ a_3, \dots, g_3 \\ a_4, \dots, g_4}} \Delta_{a_1 b_1 c_1 d_1 e_1 f_1 g_1}^{a_1 b_4 c_1 d_2 e_4 f_2 g_4} \Delta_{a_2 b_2 c_2 d_2 e_2 f_2 g_2}^{a_2 b_2 c_3 d_3 e_1 f_1 g_3} \Delta_{a_3 b_3 c_3 d_3 e_2 f_4 g_2}^{a_3 b_3 c_2 d_4 e_2 f_4 g_2} \Delta_{a_4 b_4 c_4 d_4 e_3 f_3 g_1}^{a_4 b_1 c_4 d_1 e_3 f_3 g_1} \\ &= \sum_{\substack{b_1, d_1, \dots, g_1 \\ c_2, \dots, g_2 \\ c_3, \dots, g_3 \\ b_4, d_4, \dots, g_4}} [(\text{Tr}_{\mathcal{A} \otimes \mathcal{C}} \Delta)^{\Gamma_{\mathcal{E}}}]_{b_1 d_1 e_4 f_1 g_1}^{b_4 d_2 e_1 f_2 g_4} [(\text{Tr}_{\mathcal{A} \otimes \mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}}]_{c_2 d_2 e_1 f_2 g_2}^{c_3 d_3 e_2 f_1 g_3} \\ &\quad \times [(\text{Tr}_{\mathcal{A} \otimes \mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}}]_{c_3 d_3 e_2 f_3 g_3}^{c_2 d_4 e_3 f_4 g_2} [(\text{Tr}_{\mathcal{A} \otimes \mathcal{C}} \Delta)^{\Gamma_{\mathcal{E}}}]_{b_4 d_4 e_3 f_4 g_4}^{b_1 d_1 e_4 f_3 g_1}, \end{aligned}$$

where $\Gamma_{\mathcal{E}}$ denotes the partial transposition on \mathcal{E} .

We can rewrite this using the maximally entangled $\Phi_{\mathcal{F} \otimes \mathcal{F}} = \sum_{ff'} |ff\rangle\langle f'f'|$:

Letting $\mathcal{J} := \mathcal{C} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{G}$, $P := (\text{Tr}_{\mathcal{A} \otimes \mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}}$ and $R := (P \otimes \mathbb{1}_{\mathcal{F}})(\mathbb{1}_{\mathcal{J}} \otimes \Phi_{\mathcal{F} \otimes \mathcal{F}})(P \otimes \mathbb{1}_{\mathcal{F}})$, we notice that, for all $j, j', f, f', \tilde{f}, \tilde{f}'$:

$$R_{j, f, \tilde{f}}^{j', f', \tilde{f}'} = \sum_{\substack{j'', j''' \\ f'', f''' \\ \tilde{f}'', \tilde{f}'''}} \left(P_{j, f}^{j'', f''} \delta_{\tilde{f}'' = \tilde{f}} \right) \left(\delta_{j''' = j''} \delta_{\tilde{f}''' = f''} \delta_{\tilde{f}''' = f'''} \right) \left(P_{j'', f''}^{j', f'} \delta_{\tilde{f}''' = \tilde{f}'} \right) = \sum_{j''} P_{j, f}^{j'', \tilde{f}} P_{j'', \tilde{f}}^{j', f'}.$$

Likewise, letting $\mathcal{K} := \mathcal{B} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{G}$, $Q := (\text{Tr}_{\mathcal{A} \otimes \mathcal{C}} \Delta)^{\Gamma_{\mathcal{E}}}$ and $S := (Q \otimes \mathbb{1}_{\mathcal{F}})(\mathbb{1}_{\mathcal{K}} \otimes \Phi_{\mathcal{F} \otimes \mathcal{F}})(Q \otimes \mathbb{1}_{\mathcal{F}})$, we have for all $k, k', f, f', \tilde{f}, \tilde{f}'$:

$$S_{k, f', \tilde{f}'}^{k', f, \tilde{f}} = \sum_{k''} Q_{k, f'}^{k'', \tilde{f}'} Q_{k'', \tilde{f}'}^{k', f}.$$

We now just have to make the following identifications:

- $j := (c_2, d_2, e_1, g_2)$, $j' := (c_2, d_4, e_3, g_2)$, $j'' := (c_3, d_3, e_2, g_3)$,
- $k := (b_4, d_4, e_3, g_4)$, $k' := (b_4, d_2, e_1, g_4)$, $k'' := (b_1, d_1, e_4, g_1)$,
- $f := f_2$, $f' := f_4$, $\tilde{f} := f_1$, $\tilde{f}' := f_3$,

and to notice that we can actually sum over j'' and k'' independently. We thus get:

$$\begin{aligned} \text{Tr } \Delta^{\otimes 4}(U_{\sigma_{\mathcal{A}}} \otimes \cdots \otimes U_{\sigma_{\mathcal{G}}}) &= \sum_{\substack{e_1, f_1 \\ c_2, d_2, f_2, g_2 \\ e_3, f_3 \\ b_4, d_4, f_4, g_4}} R_{c_2, d_2, e_1, g_2, f_2, f_1}^{c_2, d_4, e_3, g_2, f_4, f_3} S_{b_4, d_4, e_3, g_4, f_4, f_3}^{b_4, d_2, e_1, g_4, f_2, f_1} \\ &= \sum_{\substack{e_1, f_1 \\ d_2, f_2 \\ e_3, f_3 \\ d_4, f_4}} (\text{Tr}_{\mathcal{C} \otimes \mathcal{G}} R)_{d_2, e_1, f_2, f_1}^{d_4, e_3, f_4, f_3} (\text{Tr}_{\mathcal{B} \otimes \mathcal{G}} S)_{d_4, e_3, f_4, f_3}^{d_2, e_1, f_2, f_1} \\ &= \text{Tr}_{\mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{F}} (\text{Tr}_{\mathcal{C} \otimes \mathcal{G}} R) (\text{Tr}_{\mathcal{B} \otimes \mathcal{G}} S). \end{aligned}$$

Defining $\tilde{P} := (P \otimes \mathbf{1}_{\mathcal{F}})(\mathbf{1}_{\mathcal{J}} \otimes \sum_f |ff\rangle)$ and $\tilde{Q} := (Q \otimes \mathbf{1}_{\mathcal{F}})(\mathbf{1}_{\mathcal{J}} \otimes \sum_f |ff\rangle)$, we see that $R = \tilde{P}\tilde{P}^\dagger$ and $S = \tilde{Q}\tilde{Q}^\dagger$. Hence R and S are positive semidefinite, and so are $\text{Tr}_{\mathcal{C} \otimes \mathcal{G}} R$ and $\text{Tr}_{\mathcal{B} \otimes \mathcal{G}} S$. Thus, using the fact that, for positive semidefinite V and W , $\text{Tr} VW \leq (\text{Tr} V)(\text{Tr} W)$, we obtain

$$\text{Tr}_{\mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{F}} [(\text{Tr}_{\mathcal{C} \otimes \mathcal{G}} R) (\text{Tr}_{\mathcal{B} \otimes \mathcal{G}} S)] \leq (\text{Tr}_{\mathcal{C} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{F} \otimes \mathcal{G}} R) (\text{Tr}_{\mathcal{B} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{F} \otimes \mathcal{G}} S).$$

On right hand side,

$$\begin{aligned} \text{Tr} R &= \text{Tr}_{\mathcal{C} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}} P^2 \\ &= \text{Tr}_{\mathcal{C} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}} ((\text{Tr}_{\mathcal{A} \otimes \mathcal{B}} \Delta)^{\Gamma_{\mathcal{E}}})^2 \\ &= \text{Tr}_{\mathcal{C} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}} (\text{Tr}_{\mathcal{A} \otimes \mathcal{B}} \Delta)^2, \end{aligned}$$

and likewise, $\text{Tr} S = \text{Tr}_{\mathcal{B} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}} (\text{Tr}_{\mathcal{A} \otimes \mathcal{C}} \Delta)^2$. So, we eventually arrive at

$$\text{Tr} \Delta^{\otimes 4} (U_{\sigma_{\mathcal{A}}} \otimes \cdots \otimes U_{\sigma_{\mathcal{G}}}) \leq \left[\text{Tr}_{\mathcal{C} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}} (\text{Tr}_{\mathcal{A} \otimes \mathcal{B}} \Delta)^2 \right] \left[\text{Tr}_{\mathcal{B} \otimes \mathcal{D} \otimes \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}} (\text{Tr}_{\mathcal{A} \otimes \mathcal{C}} \Delta)^2 \right]. \quad (\text{A2})$$

With this inequality as a tool, we can now return to our initial problem: For all $\underline{\pi} \in \mathfrak{A}^K = \{\text{id}, (14), (23), (1234), (1432), (12)(34), (14)(23)\}^K$, we can define the following factors of the global Hilbert space \mathcal{H} :

$$\begin{aligned} \mathcal{A}(\underline{\pi}) &:= \bigotimes_{j \text{ s.t. } \pi_j = \text{id}} \mathcal{H}_j, & \mathcal{B}(\underline{\pi}) &:= \bigotimes_{j \text{ s.t. } \pi_j = (14)} \mathcal{H}_j, & \mathcal{C}(\underline{\pi}) &:= \bigotimes_{j \text{ s.t. } \pi_j = (23)} \mathcal{H}_j, \\ \mathcal{D}(\underline{\pi}) &:= \bigotimes_{j \text{ s.t. } \pi_j = (1234)} \mathcal{H}_j, & \mathcal{E}(\underline{\pi}) &:= \bigotimes_{j \text{ s.t. } \pi_j = (1432)} \mathcal{H}_j, \\ \mathcal{F}(\underline{\pi}) &:= \bigotimes_{j \text{ s.t. } \pi_j = (12)(34)} \mathcal{H}_j, & \mathcal{G}(\underline{\pi}) &:= \bigotimes_{j \text{ s.t. } \pi_j = (14)(23)} \mathcal{H}_j, \end{aligned}$$

so that clearly, $\mathcal{H} = \mathcal{A}(\underline{\pi}) \otimes \mathcal{B}(\underline{\pi}) \otimes \mathcal{C}(\underline{\pi}) \otimes \mathcal{D}(\underline{\pi}) \otimes \mathcal{E}(\underline{\pi}) \otimes \mathcal{F}(\underline{\pi}) \otimes \mathcal{G}(\underline{\pi})$. Hence, using successively the two inequalities (A1) and (A2), we have:

$$\begin{aligned} \sum_{\underline{\sigma} \in \mathfrak{S}_4^K} \text{Tr} (\Delta^{\otimes 4} U_{\underline{\sigma}}) &\leq \sum_{\underline{\sigma} \in \mathfrak{S}_4^K} \left\{ \frac{1}{2} \text{Tr} (\Delta^{\otimes 4} U_{\underline{\sigma}^L}) + \frac{1}{2} \text{Tr} (\Delta^{\otimes 4} U_{\underline{\sigma}^R}) \right\} \\ &\leq \sum_{\underline{\sigma} \in \mathfrak{S}_4^K} \left\{ \frac{1}{2} \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^L) \otimes \mathcal{B}(\underline{\sigma}^L)} \Delta \right)^2 \right] \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^L) \otimes \mathcal{C}(\underline{\sigma}^L)} \Delta \right)^2 \right] \right. \\ &\quad \left. + \frac{1}{2} \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^R) \otimes \mathcal{B}(\underline{\sigma}^R)} \Delta \right)^2 \right] \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^R) \otimes \mathcal{C}(\underline{\sigma}^R)} \Delta \right)^2 \right] \right\} \\ &= \sum_{\underline{\sigma} \in \mathfrak{S}_4^K} \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^L) \otimes \mathcal{B}(\underline{\sigma}^L)} \Delta \right)^2 \right] \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^L) \otimes \mathcal{C}(\underline{\sigma}^L)} \Delta \right)^2 \right] \\ &\leq \sum_{\underline{\sigma} \in \mathfrak{S}_4^K} \left\{ \frac{1}{2} \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^L) \otimes \mathcal{B}(\underline{\sigma}^L)} \Delta \right)^2 \right]^2 + \frac{1}{2} \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^L) \otimes \mathcal{C}(\underline{\sigma}^L)} \Delta \right)^2 \right]^2 \right\} \\ &= \sum_{\underline{\sigma} \in \mathfrak{S}_4^K} \left[\text{Tr} \left(\text{Tr}_{\mathcal{A}(\underline{\sigma}^L) \otimes \mathcal{B}(\underline{\sigma}^L)} \Delta \right)^2 \right]^2, \end{aligned}$$

where in the last lines we have made use of the symmetry between $\underline{\sigma}^L$ and $\underline{\sigma}^R$ on the one hand, and that between $\mathcal{B}(\underline{\sigma}^L)$ and $\mathcal{C}(\underline{\sigma}^L)$ on the other, when $\underline{\sigma}$ ranges over \mathfrak{S}_4^K .

This concludes the main trace estimate. The rest of the argument is combinatorial. Observe that among the 24 permutations σ of \mathfrak{S}_4 , two are such that $\sigma^L = \text{id}$ [namely id and (34)], and four are such that $\sigma^L = (14)$ [namely (14) , (13) , (134) and (143)]. Hence, for all subsets $I \subset [K]$, there are $6^{|I|} \times 18^{K-|I|}$ K -tuples of permutations $\underline{\sigma}$ such that σ_i^L is either id or (14) for $i \in I$, and σ_i^L is neither id nor (14) for $i \notin I$. Therefore, we finally obtain:

$$\text{Tr} \left(\Delta^{\otimes 4} \left(\sum_{\underline{\sigma} \in \mathfrak{S}_4^K} U_{\underline{\sigma}} \right) \right) \leq 18^K \sum_{I \subset [K]} \left[\text{Tr} (\text{Tr}_I \Delta)^2 \right]^2 \leq 18^K \left[\sum_{I \subset [K]} \text{Tr} (\text{Tr}_I \Delta)^2 \right]^2,$$

which is what we wanted to prove. \square